

Kurtosis Tests for Multivariate Normality with Monotone Incomplete Data

Tomoya Yamada,^{*} Megan M. Romer,[†] and Donald St. P. Richards[‡]

November 27, 2014

Abstract

We consider the problem of testing multivariate normality when the data consists of a random sample of two-step monotone incomplete observations. We define for such data a generalization of Mardia's statistic for measuring kurtosis, derive the asymptotic non-null distribution of the statistic under certain regularity conditions and against a broad class of alternatives, and give an application to a well-known data set on cholesterol measurements.

1 Introduction

In many statistical applications, and especially in epidemiology and biostatistics, incomplete data arise for a variety of reasons; cf., Eaton and Kariya (1983), Garren and Peddada (2000), Little and Rubin (2002), Peddada, Harris, and Davidov (2010), Krishnamoorthy and Yu (2012), and Davidov and Peddada (2013). Consequently, much work has been done on explicit formulas that allow for statistical inference with incomplete data to achieve specified levels of significance.

In this paper, we consider the problem of kurtosis tests for multivariate normality using two-step monotone incomplete data. Following our earlier work on monotone incomplete multivariate normal data (Chang and Richards, 2009, 2010; Richards and Yamada, 2010; Romer, 2009; Romer and Richards, 2010, 2013; Yamada, 2013), we write the data in the form

$$\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{Y}_1 \end{pmatrix} \begin{pmatrix} \mathbf{X}_2 \\ \mathbf{Y}_2 \end{pmatrix} \cdots \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} \mathbf{Y}_{n+1} \mathbf{Y}_{n+2} \cdots \mathbf{Y}_N \quad (1.1)$$

where each \mathbf{X}_j , $1 \leq j \leq n$, is $p \times 1$ and each \mathbf{Y}_j , $1 \leq j \leq N$, is $q \times 1$.

As in our earlier work, we assume that the data are missing completely at random (MCAR). It follows from results of Eaton and Kariya (1983), Hao and Krishnamoorthy (2001), and others that an explicit solution for the likelihood equations for the covariance matrix requires the MCAR assumption. Thus, the MCAR assumption and the monotone data pattern ensure the validity of likelihood inference and a unique, explicit solution to the likelihood equations.

In this paper, we define for data of the form (1.1) a generalization of Mardia's statistic for testing kurtosis. We derive the asymptotic non-null and null distributions of the new statistic under certain regularity conditions. We apply our results to a well-known Pennsylvania cholesterol data set (Ryan, Joiner, and Cryer, 2005, p. 267) which has been used widely to illustrate statistical methods for

^{*}Department of Economics, Sapporo Gakuin University, 11 Bunkyo-dai, Ebetsu, Hokkaido, Japan. Research supported by a sabbatical leave-of-absence from Sapporo Gakuin University. E-mail address: tomoya@sgu.ac.jp

[†]Department of Statistics, Penn State University, University Park, PA 16802, U.S.A. Research supported in part by National Science Foundation grant DMS-1309808. E-mail address: mmd198@stat.psu.edu

[‡]Department of Statistics, Penn State University, University Park, PA 16802, U.S.A. Research supported in part by National Science Foundation grant DMS-1309808. E-mail address: richards@stat.psu.edu

2000 Mathematics Subject Classification: Primary 62C15, 62H10; Secondary 60D10, 62E15.

Key words and phrases: Asymptotic expansion; Mardia's statistic; maximum likelihood estimation; missing data; monotone incomplete data; regression imputation; testing for multivariate normality.

analyzing monotone incomplete multivariate data. Our results provide an invariant test of normality for that data, adding to the literature on that subject (see Henze (2002) for an extensive account of invariant testing procedures in the complete case), and we complement results of Romer (2009) and Romer and Richards (2013), where exploratory methods were applied to the cholesterol data set.

Our results are as follows. In Section 2, we provide the background for testing kurtosis with the data (1.1). We define in Section 3 the kurtosis statistic, $b_{2,p,q}$, and prove that $b_{2,p,q}$ is identical to a statistic constructed from the observed data and from data imputed by linear regression methods. Further, we establish an invariance property of $b_{2,p,q}$ which allows us to reduce the general problem to the canonical case in which the population has mean zero and identity covariance matrix. We state in Section 4 the null and non-null asymptotic distributions of $b_{2,p,q}$ corresponding, respectively to the normal case and to a broad class of alternatives defined by moment conditions on the distribution, and we apply those results to the Pennsylvania cholesterol data, reaching conclusions similar to Romer (2009). We derive in Appendix A asymptotic expansions of $\hat{\Sigma}$ and $\hat{\Sigma}^{-1}$, where $\hat{\Sigma}$ is an estimator of Σ , the covariance matrix of the population underlying the sample (1.1), and then in Appendix B we apply those expansions to derive asymptotic expansions of two statistics used to construct $b_{2,p,q}$. Finally, we obtain in Appendix C the null and non-null asymptotic distributions of $b_{2,p,q}$.

2 Notation and preliminaries

Throughout the paper, we follow the notation of Chang and Richards (2009, 2010). Thus, all matrices and vectors are written in boldface type; \mathbf{I}_d denotes the identity matrix of order d ; and $\mathbf{0}$ denotes any matrix or vector of zeros, the dimension of which will be clear from the context. We also let $\tau = n/N$ denote the proportion of observations in (1.1) that are complete, and set $\bar{\tau} = 1 - \tau \equiv (N - n)/N$.

Define the sample means

$$\bar{\mathbf{X}} = \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j, \quad \bar{\mathbf{Y}}_1 = \frac{1}{n} \sum_{j=1}^n \mathbf{Y}_j, \quad \bar{\mathbf{Y}}_2 = \frac{1}{N-n} \sum_{j=n+1}^N \mathbf{Y}_j, \quad (2.1)$$

and

$$\bar{\mathbf{Y}} = \frac{1}{N} \sum_{j=1}^N \mathbf{Y}_j \equiv \tau \bar{\mathbf{Y}}_1 + \bar{\tau} \bar{\mathbf{Y}}_2. \quad (2.2)$$

We also define corresponding matrices of sums of squares and products,

$$\begin{aligned} \mathbf{A}_{11} &= \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})', & \mathbf{A}_{12} &= \mathbf{A}_{21}' = \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{Y}_j - \bar{\mathbf{Y}}_1)', \\ \mathbf{A}_{22,n} &= \sum_{j=1}^n (\mathbf{Y}_j - \bar{\mathbf{Y}}_1)(\mathbf{Y}_j - \bar{\mathbf{Y}}_1)', & \mathbf{A}_{22,N} &= \sum_{j=1}^N (\mathbf{Y}_j - \bar{\mathbf{Y}})(\mathbf{Y}_j - \bar{\mathbf{Y}})', \end{aligned} \quad (2.3)$$

set $\mathbf{A}_{11 \cdot 2, n} = \mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22, n}^{-1} \mathbf{A}_{21}$, and let

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22, n} \end{pmatrix}. \quad (2.4)$$

Let $\boldsymbol{\mu}$ denote the mean and Σ the covariance matrix of the population underlying the data (1.1); we assume that Σ is nonsingular. We partition $\boldsymbol{\mu}$ and Σ similar to (1.1), so that

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are $p \times 1$ and $q \times 1$ vectors, respectively, and the submatrices $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}'$, and $\boldsymbol{\Sigma}_{22}$ are of order $p \times p$, $p \times q$, and $q \times q$, respectively. We also define

$$\hat{\boldsymbol{\mu}} = \begin{pmatrix} \hat{\boldsymbol{\mu}}_1 \\ \hat{\boldsymbol{\mu}}_2 \end{pmatrix}, \quad \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} \hat{\boldsymbol{\Sigma}}_{11} & \hat{\boldsymbol{\Sigma}}_{12} \\ \hat{\boldsymbol{\Sigma}}_{21} & \hat{\boldsymbol{\Sigma}}_{22} \end{pmatrix},$$

where

$$\hat{\boldsymbol{\mu}}_1 = \bar{\mathbf{X}} - \bar{\tau} \mathbf{A}_{12} \mathbf{A}_{22,n}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2), \quad \hat{\boldsymbol{\mu}}_2 = \bar{\mathbf{Y}}, \quad (2.5)$$

and

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_{11} &= \frac{1}{n} \mathbf{A}_{11 \cdot 2, n} + \frac{1}{N} \mathbf{A}_{12} \mathbf{A}_{22, n}^{-1} \mathbf{A}_{22, N} \mathbf{A}_{22, n}^{-1} \mathbf{A}_{21}, \\ \hat{\boldsymbol{\Sigma}}_{12} &= \hat{\boldsymbol{\Sigma}}_{21}' = \frac{1}{N} \mathbf{A}_{12} \mathbf{A}_{22, n}^{-1} \mathbf{A}_{22, N}, \quad \hat{\boldsymbol{\Sigma}}_{22} = \frac{1}{N} \mathbf{A}_{22, N}. \end{aligned} \quad (2.6)$$

If the underlying population is multivariate normal, denoted by $N_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are the maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, respectively. We refer to Chang and Richards (2009, 2010), Richards and Yamada (2010), Romer (2009), and Romer and Richards (2010) for results on the distributions of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, and inference for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

As noted by Chang and Richards (2009, p. 1886), the statistical model underlying two-step monotone incomplete multivariate normal data is related to double sampling designs, in which additional data are collected on a subset of variables in order to improve estimation of a parameter; cf. Little (1976, p. 594) and Cohn, Davidov, and Haitovsky (2008).

3 Testing kurtosis with monotone incomplete data

3.1 The kurtosis statistic

Consider a d -dimensional multivariate population represented by a random vector \mathbf{Z} with mean vector $\boldsymbol{\mu}$ and nonsingular covariance matrix $\boldsymbol{\Sigma}$, and kurtosis parameter

$$\beta_{2,d} = E[(\mathbf{Z} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{Z} - \boldsymbol{\mu})]^2.$$

To perform inference for $\beta_{2,d}$ with the monotone incomplete data (1.1), we define the statistic

$$\begin{aligned} b_{2,p,q} &= \frac{1}{N} \left\{ c_1 \sum_{j=1}^n \left[\left(\begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} - \hat{\boldsymbol{\mu}} \right)' \hat{\boldsymbol{\Sigma}}^{-1} \left(\begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} - \hat{\boldsymbol{\mu}} \right) \right]^2 \right. \\ &\quad \left. + c_2 \sum_{j=n+1}^N \left[(\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_2)' \hat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_2) \right]^2 \right\}, \end{aligned} \quad (3.1)$$

where $c_1, c_2 > 0$ are constants. In general, c_1 and c_2 can depend on n and N subject to the conditions $c_1 = O(\tau)$, $c_2 = O(\bar{\tau})$, and $c_1, c_2 \not\rightarrow 0$ as $n, N \rightarrow \infty$; e.g., $(c_1, c_2) = (\tau, \bar{\tau})$ with $n/N \rightarrow \delta \in (0, 1)$. Alternatively, (c_1, c_2) can be chosen to minimize σ^2 , the asymptotic variance of $b_{2,p,q}$ under the null hypothesis; this can be obtained by minimizing, over all (c_1, c_2) subject to a suitable constraint on c_1 and c_2 , a formula for σ^2 in (4.12).

Each term in (3.1) is an analog of the well-known statistic of Mardia (1970, 1974) for testing kurtosis with complete data, and our usage of different weights is due to the fact that the incomplete data \mathbf{Y}_j , $j = n+1, \dots, N$ provide partial information about the population.

We may also motivate the statistic $b_{2,p,q}$ as follows: First, we impute each missing observation \mathbf{X}_j , $j = n+1, \dots, N$, using a linear regression imputation scheme,

$$\widehat{\mathbf{X}}_j = \widehat{E}(\mathbf{X}_j | \mathbf{Y}_j) \equiv \hat{\boldsymbol{\mu}}_1 + \hat{\boldsymbol{\Sigma}}_{12} \hat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \hat{\boldsymbol{\mu}}_2), \quad (3.2)$$

which is motivated by the formula for the conditional expectation of a partitioned multivariate normally distributed random vector. Under the hypothesis of multivariate normality, $\widehat{\mathbf{X}}_j$ is the maximum likelihood estimator of $E(\mathbf{X}_j|\mathbf{Y}_j)$, the conditional expectation of \mathbf{X}_j given \mathbf{Y}_j . Second, we use as our data the merged sets of observed and imputed data vectors,

$$\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{Y}_1 \end{pmatrix} \begin{pmatrix} \mathbf{X}_2 \\ \mathbf{Y}_2 \end{pmatrix} \cdots \begin{pmatrix} \mathbf{X}_n \\ \mathbf{Y}_n \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{X}}_{n+1} \\ \mathbf{Y}_{n+1} \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{X}}_{n+2} \\ \mathbf{Y}_{n+2} \end{pmatrix} \cdots \begin{pmatrix} \widehat{\mathbf{X}}_N \\ \mathbf{Y}_N \end{pmatrix}. \quad (3.3)$$

To perform inference about $\beta_{2,d}$, it is natural to use the statistic

$$\begin{aligned} \widehat{b}_{2,p,q} = \frac{1}{N} \Big\{ & c_1 \sum_{j=1}^n \left[\left(\begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} - \widehat{\boldsymbol{\mu}} \right)' \widehat{\boldsymbol{\Sigma}}^{-1} \left(\begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} - \widehat{\boldsymbol{\mu}} \right) \right]^2 \\ & + c_2 \sum_{j=n+1}^N \left[\left(\begin{pmatrix} \widehat{\mathbf{X}}_j \\ \mathbf{Y}_j \end{pmatrix} - \widehat{\boldsymbol{\mu}} \right)' \widehat{\boldsymbol{\Sigma}}^{-1} \left(\begin{pmatrix} \widehat{\mathbf{X}}_j \\ \mathbf{Y}_j \end{pmatrix} - \widehat{\boldsymbol{\mu}} \right) \right]^2 \Big\}, \end{aligned} \quad (3.4)$$

an analog of Mardia's statistic based on the vectors in (3.3). We again use possibly different weights, c_1 and c_2 , to reflect the fact that some data are imputed, hence they provide less information about $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ than in the case in which all observations are fully observed. It is remarkable that $b_{2,p,q} \equiv \widehat{b}_{2,p,q}$, a result established in Theorem 3.1 given in Section 3.2.

Set $\mathbf{Z}_j = \begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix}$, $j = 1, \dots, n$; then, (3.1) becomes

$$b_{2,p,q} = \frac{1}{N} \left\{ c_1 \sum_{j=1}^n [(\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})]^2 + c_2 \sum_{j=n+1}^N [(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)' \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)]^2 \right\}.$$

Also, let

$$\bar{\mathbf{Z}}_1 = \frac{1}{n} \sum_{j=1}^n \mathbf{Z}_j \equiv \begin{pmatrix} \bar{\mathbf{X}}_1 \\ \bar{\mathbf{Y}}_1 \end{pmatrix}, \quad (3.5)$$

and

$$\tilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{A}_{12} \mathbf{A}_{22,n}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix} \equiv \mathbf{A} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}. \quad (3.6)$$

By a direct calculation using (2.4) and (2.5), we deduce that

$$\mathbf{Z}_j - \widehat{\boldsymbol{\mu}} = \mathbf{Z}_j - \bar{\mathbf{Z}}_1 + \bar{\mathbf{Z}}_1 - \widehat{\boldsymbol{\mu}}. \quad (3.7)$$

3.2 An invariance property of $b_{2,p,q}$

Define the statistics

$$b_{2,p,q}^{(1)} = \sum_{j=1}^n [(\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})]^2,$$

and

$$b_{2,p,q}^{(2)} = \sum_{j=n+1}^N [(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)' \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)]^2.$$

Then,

$$b_{2,p,q} = \frac{1}{N} \left(c_1 b_{2,p,q}^{(1)} + c_2 b_{2,p,q}^{(2)} \right). \quad (3.8)$$

Let $\mathbf{\Lambda}_{11}$ and $\mathbf{\Lambda}_{22}$ be $p \times p$ and $q \times q$ positive definite matrices, respectively; let $\mathbf{\Lambda}_{12}$ be a $p \times q$ matrix; and let $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ be $p \times 1$ and $q \times 1$ vectors, respectively. Set

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{22} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{I}_p & \mathbf{\Lambda}_{12} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix}, \quad \boldsymbol{\nu} = \begin{pmatrix} \boldsymbol{\nu}_1 \\ \boldsymbol{\nu}_2 \end{pmatrix},$$

and consider the group of affine transformations of the data (1.1) of the form

$$\begin{aligned} \begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} &\rightarrow \mathbf{\Lambda} \mathbf{C} \begin{pmatrix} \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix} + \boldsymbol{\nu}, \quad j = 1, \dots, n \\ \mathbf{Y}_j &\rightarrow \mathbf{\Lambda}_{22} \mathbf{Y}_j + \boldsymbol{\nu}_2, \quad j = n+1, \dots, N \end{aligned} \tag{3.9}$$

Now we have the following result:

Theorem 3.1. *The kurtosis statistics $b_{2,p,q}^{(1)}$, $b_{2,p,q}^{(2)}$, and $b_{2,p,q}$ are invariant under the transformations (3.9). Moreover, $b_{2,p,q} \equiv \widehat{b}_{2,p,q}$.*

Proof. Under the transformation (3.9), we verify using (2.1) and (2.2) that $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}_1$, $\bar{\mathbf{Y}}_2$, and $\bar{\mathbf{Y}}$ are transformed to $\mathbf{\Lambda}_{11}(\bar{\mathbf{X}}_1 + \mathbf{\Lambda}_{12}\bar{\mathbf{Y}}_1) + \boldsymbol{\nu}_1$, $\mathbf{\Lambda}_{22}\bar{\mathbf{Y}}_1 + \boldsymbol{\nu}_2$, $\mathbf{\Lambda}_{22}\bar{\mathbf{Y}}_2 + \boldsymbol{\nu}_2$, and $\mathbf{\Lambda}_{22}\bar{\mathbf{Y}} + \boldsymbol{\nu}_2$, respectively. Further, by (2.3), the matrix \mathbf{A} in (2.4) is transformed to $\mathbf{\Lambda} \mathbf{C} \mathbf{A} \mathbf{C}' \mathbf{\Lambda}$, i.e.,

$$\begin{aligned} \mathbf{A}_{11} &\rightarrow \mathbf{\Lambda}_{11}(\mathbf{A}_{11} + \mathbf{\Lambda}_{12}\mathbf{A}_{21} + \mathbf{A}_{12}\mathbf{\Lambda}_{21} + \mathbf{\Lambda}_{12}\mathbf{A}_{22,n}\mathbf{\Lambda}_{21})\mathbf{\Lambda}_{11} \\ \mathbf{A}_{12} &\rightarrow \mathbf{\Lambda}_{11}(\mathbf{A}_{12} + \mathbf{\Lambda}_{12}\mathbf{A}_{22,n})\mathbf{\Lambda}_{22}, \\ \mathbf{A}_{22,n} &\rightarrow \mathbf{\Lambda}_{22}\mathbf{A}_{22,n}\mathbf{\Lambda}_{22} \end{aligned} \tag{3.10}$$

Hence $\mathbf{A}_{11 \cdot 2, n} \rightarrow \mathbf{\Lambda}_{11}\mathbf{A}_{11 \cdot 2, n}\mathbf{\Lambda}_{11}$ and $\mathbf{A}_{22, N} \rightarrow \mathbf{\Lambda}_{22}\mathbf{A}_{22, N}\mathbf{\Lambda}_{22}$. Further, it follows from (2.6) and (3.10) that $\widehat{\boldsymbol{\Sigma}}_{11 \cdot 2}$ and $\widehat{\boldsymbol{\Sigma}}_{22}$ are transformed to $\mathbf{\Lambda}_{11}\widehat{\boldsymbol{\Sigma}}_{11 \cdot 2}\mathbf{\Lambda}_{11}$ and $\mathbf{\Lambda}_{22}\widehat{\boldsymbol{\Sigma}}_{22}\mathbf{\Lambda}_{22}$, respectively.

By a well-known quadratic identity (Anderson, 2003, p. 63, Exercise 2.54), for $j = 1, \dots, n$,

$$\begin{aligned} &(\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}}) \\ &\equiv \begin{pmatrix} \mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 \\ \mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2 \end{pmatrix}' \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_{11} & \widehat{\boldsymbol{\Sigma}}_{12} \\ \widehat{\boldsymbol{\Sigma}}_{21} & \widehat{\boldsymbol{\Sigma}}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 \\ \mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} \\ &= (\mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2))' \widehat{\boldsymbol{\Sigma}}_{11 \cdot 2}^{-1} (\mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)) \\ &\quad + (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)' \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2). \end{aligned} \tag{3.11}$$

It follows from (2.5) that $\mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2) = \mathbf{X}_j - \bar{\mathbf{X}} - \mathbf{A}_{12}\mathbf{A}_{22,n}^{-1}(\mathbf{Y}_j - \bar{\mathbf{Y}}_1)$, and the latter expression is transformed by (3.9) to

$$\begin{aligned} &\mathbf{\Lambda}_{11}(\mathbf{X}_j + \mathbf{\Lambda}_{12}\mathbf{Y}_j) - \mathbf{\Lambda}_{11}(\bar{\mathbf{X}} + \mathbf{\Lambda}_{12}\bar{\mathbf{Y}}_1) - \mathbf{\Lambda}_{11}(\mathbf{A}_{12} + \mathbf{\Lambda}_{12}\mathbf{A}_{22,n})\mathbf{A}_{22,n}^{-1}(\mathbf{Y}_j - \bar{\mathbf{Y}}_1) \\ &= \mathbf{\Lambda}_{11}(\mathbf{X}_j - \bar{\mathbf{X}} - \mathbf{A}_{12}\mathbf{A}_{22,n}^{-1}(\mathbf{Y}_j - \bar{\mathbf{Y}}_1)) \\ &\equiv \mathbf{\Lambda}_{11}(\mathbf{X}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12}\widehat{\boldsymbol{\Sigma}}_{22}^{-1}(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)). \end{aligned}$$

Recall that $\widehat{\boldsymbol{\Sigma}}_{11 \cdot 2}$ is transformed to $\mathbf{\Lambda}_{11}\widehat{\boldsymbol{\Sigma}}_{11 \cdot 2}\mathbf{\Lambda}_{11}$, so it follows that the first term in (3.11) remains invariant under (3.9).

In similar fashion, (3.9) transforms $\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2$ to $\mathbf{\Lambda}_{22}(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)$, $j = 1, \dots, N$ and, as noted earlier, also transforms $\widehat{\boldsymbol{\Sigma}}_{22}$ to $\mathbf{\Lambda}_{22}\widehat{\boldsymbol{\Sigma}}_{22}\mathbf{\Lambda}_{22}$; consequently, the second term in (3.11) remains invariant under (3.9). It follows that $b_{2,p,q}^{(1)}$ and $b_{2,p,q}^{(2)}$ each are invariant under (3.9), so is $b_{2,p,q}$.

Finally, to show that $b_{2,p,q} \equiv \widehat{b}_{2,p,q}$, we apply the earlier quadratic identity (Anderson, 2003, loc. cit.) for $j = n+1, \dots, N$ to obtain

$$\begin{aligned} & \begin{pmatrix} \widehat{\mathbf{X}}_j - \widehat{\boldsymbol{\mu}}_1 \\ \mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2 \end{pmatrix}' \begin{pmatrix} \widehat{\boldsymbol{\Sigma}}_{11} & \widehat{\boldsymbol{\Sigma}}_{12} \\ \widehat{\boldsymbol{\Sigma}}_{21} & \widehat{\boldsymbol{\Sigma}}_{22} \end{pmatrix}^{-1} \begin{pmatrix} \widehat{\mathbf{X}}_j - \widehat{\boldsymbol{\mu}}_1 \\ \mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2 \end{pmatrix} \\ &= (\widehat{\mathbf{X}}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12} \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2))' \widehat{\boldsymbol{\Sigma}}_{11.2}^{-1} (\widehat{\mathbf{X}}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12} \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)) \\ & \quad + (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)' \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2). \end{aligned}$$

By the definition in (3.2) of $\widehat{\mathbf{X}}_j$, it follows that $\widehat{\mathbf{X}}_j - \widehat{\boldsymbol{\mu}}_1 - \widehat{\boldsymbol{\Sigma}}_{12} \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2) \equiv \mathbf{0}$. Therefore the last terms in (3.1) and (3.4) are identical, so we obtain $b_{2,p,q} \equiv \widehat{b}_{2,p,q}$. \square

Remark 3.2. In the multivariate normal case, Romer and Richards (2010, Proposition 2.1) showed that the invariance of $b_{2,p,q}$ is due to the fact that $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$, being maximum likelihood estimators, are equivariant. Specifically, $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ are transformed under (3.9) to $\boldsymbol{\Lambda} \mathbf{C} \widehat{\boldsymbol{\mu}} + \boldsymbol{\nu}$ and $\boldsymbol{\Lambda} \mathbf{C} \widehat{\boldsymbol{\Sigma}} \mathbf{C}' \boldsymbol{\Lambda}$, respectively; therefore the quadratic form $(\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})$, $j = 1, \dots, n$ is transformed to

$$\begin{aligned} & ((\boldsymbol{\Lambda} \mathbf{C} \mathbf{Z}_j + \boldsymbol{\nu}) - (\boldsymbol{\Lambda} \mathbf{C} \widehat{\boldsymbol{\mu}} + \boldsymbol{\nu}))' (\boldsymbol{\Lambda} \mathbf{C} \widehat{\boldsymbol{\Sigma}} \mathbf{C}' \boldsymbol{\Lambda})^{-1} ((\boldsymbol{\Lambda} \mathbf{C} \mathbf{Z}_j + \boldsymbol{\nu}) - (\boldsymbol{\Lambda} \mathbf{C} \widehat{\boldsymbol{\mu}} + \boldsymbol{\nu})) \\ & \equiv (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}})' \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \widehat{\boldsymbol{\mu}}), \end{aligned}$$

and this proves that $b_{2,p,q}^{(1)}$ is invariant under the group (3.9). It can be shown similarly that each quadratic form $(\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)' \widehat{\boldsymbol{\Sigma}}_{22}^{-1} (\mathbf{Y}_j - \widehat{\boldsymbol{\mu}}_2)$, $j = n+1, \dots, N$ is invariant under (3.9), hence so is $b_{2,p,q}^{(2)}$. Therefore, in the multivariate normal case, the invariance of $b_{2,p,q}$ under (3.9) is a consequence of equivariance.

In non-normal cases, however, the detailed computations in the proof of Theorem 3.1 are necessary to prove that $b_{2,p,q}$ is invariant under (3.9). Moreover, now that it has been established that $b_{2,p,q}$ is invariant under (3.9), we then choose $\boldsymbol{\Lambda}_{11} = \boldsymbol{\Sigma}_{11.2}^{-1/2}$, $\boldsymbol{\Lambda}_{22} = \boldsymbol{\Sigma}_{22}^{-1/2}$, $\boldsymbol{\Lambda}_{12} = -\boldsymbol{\Sigma}_{12} \boldsymbol{\Sigma}_{22}^{-1}$, and $\boldsymbol{\nu} = -\boldsymbol{\Lambda} \mathbf{C} \boldsymbol{\mu}$ to reduce the data to being mutually independent and monotone incomplete with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Lambda} \mathbf{C} \boldsymbol{\Sigma} \mathbf{C}' \boldsymbol{\Lambda}' = \mathbf{I}_{p+q}$. Therefore, in deriving the distribution of $b_{2,p,q}$, we assume without loss of generality that $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}_{p+q}$.

4 The null and non-null asymptotic distributions of $b_{2,p,q}$

For the $(p+q)$ -dimensional random vector $\mathbf{Z} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}$, define

$$\boldsymbol{\Xi} = E((\mathbf{Z} \mathbf{Z}')^2) = E(\mathbf{Z} (\mathbf{Z}' \mathbf{Z}) \mathbf{Z}') = E(\|\mathbf{Z}\|^2 \mathbf{Z} \mathbf{Z}'), \quad (4.1)$$

and write $\boldsymbol{\Xi}$ in partitioned form,

$$\boldsymbol{\Xi} \equiv \begin{pmatrix} \boldsymbol{\Xi}_{11} & \boldsymbol{\Xi}_{12} \\ \boldsymbol{\Xi}_{21} & \boldsymbol{\Xi}_{22} \end{pmatrix} = \begin{pmatrix} E(\|\mathbf{Z}\|^2 \mathbf{X} \mathbf{X}') & E(\|\mathbf{Z}\|^2 \mathbf{X} \mathbf{Y}') \\ E(\|\mathbf{Z}\|^2 \mathbf{Y} \mathbf{X}') & E(\|\mathbf{Z}\|^2 \mathbf{Y} \mathbf{Y}') \end{pmatrix}. \quad (4.2)$$

Define

$$\boldsymbol{\Xi}^* = E((\mathbf{Y} \mathbf{Y}')^2), \quad \boldsymbol{\Theta}^* = E(\|\mathbf{Y}\|^2 \mathbf{Y}); \quad (4.3)$$

and set

$$\widetilde{\boldsymbol{\Xi}} \equiv \begin{pmatrix} \widetilde{\boldsymbol{\Xi}}_{11} & \widetilde{\boldsymbol{\Xi}}_{12} \\ \widetilde{\boldsymbol{\Xi}}_{21} & \widetilde{\boldsymbol{\Xi}}_{22} \end{pmatrix} = \begin{pmatrix} c_1 \boldsymbol{\Xi}_{11} & c_1 \boldsymbol{\Xi}_{12} \\ c_1 \boldsymbol{\Xi}_{21} & c_1 \tau \boldsymbol{\Xi}_{22} + c_2 \bar{\tau} \boldsymbol{\Xi}^* \end{pmatrix}, \quad (4.4)$$

$$\Theta \equiv \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix} = \begin{pmatrix} E(\|Z\|^2 X) \\ E(\|Z\|^2 Y) \end{pmatrix}, \quad (4.5)$$

and

$$\tilde{\Theta} \equiv \begin{pmatrix} \tilde{\Theta}_1 \\ \tilde{\Theta}_2 \end{pmatrix} = \begin{pmatrix} c_1 \Theta_1 \\ c_1 \tau \Theta_2 + c_2 \bar{\tau} \Theta^* \end{pmatrix}. \quad (4.6)$$

Our main result, given in the following theorem, provides the non-null distribution of $b_{2,p,q}$ for a large class of alternatives.

Theorem 4.1. *Suppose that the monotone incomplete random sample (1.1) is drawn from a population modeled by a random vector $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ such that $E(Z) = \mathbf{0}$, $\text{Cov}(Z) = I_{p+q}$, and $E\|Z\|^8 < \infty$. For $n, N \rightarrow \infty$ with $n/N \rightarrow \delta \in (0, 1)$, we have*

$$N^{1/2} (b_{2,p,q} - \nu) / \sigma \xrightarrow{L} N(0, 1), \quad (4.7)$$

where

$$\nu = c_1 \tau E(\|Z\|^4) + c_2 \bar{\tau} E(\|Y\|^4) \quad (4.8)$$

and

$$\begin{aligned} \sigma^2 = & \tau \left\{ c_1^2 \text{Var}\|Z\|^4 + 4 \text{Var}(Z' \tilde{\Xi} Z) + 16 \tilde{\Theta}' \tilde{\Theta} \right. \\ & \left. - 4 c_1 \text{Cov}(\|Z\|^4, Z' \tilde{\Xi} Z) - 8 c_1 E\|Z\|^4 Z' \tilde{\Theta} + 16 E[Z' \tilde{\Xi} Z Z'] \tilde{\Theta} \right\} \\ & + \bar{\tau} \left\{ c_2^2 \text{Var}\|Y\|^4 + 4 \text{Var}(Y' \tilde{\Xi}_{22} Y) + 16 \tilde{\Theta}_2' \tilde{\Theta}_2 \right. \\ & \left. - 4 c_2 \text{Cov}(\|Y\|^4, Y' \tilde{\Xi}_{22} Y) - 8 c_2 E\|Y\|^4 Y' \tilde{\Theta}_2 + 16 E[Y' \tilde{\Xi}_{22} Y Y'] \tilde{\Theta}_2 \right\}. \end{aligned} \quad (4.9)$$

The proof of this result is provided in Appendix C. For the case in which $Z \sim N_{p+q}(\mathbf{0}, I_{p+q})$, the limiting distribution (4.7) reduces to the following result on the null distribution of $b_{2,p,q}$.

Corollary 4.2. *Suppose that the monotone incomplete sample (1.1) is drawn from $N_{p+q}(\mathbf{0}, I_{p+q})$. For $n, N \rightarrow \infty$ with $n/N \rightarrow \delta \in (0, 1)$, we have*

$$N^{1/2} (b_{2,p,q} - \nu) / \sigma \xrightarrow{L} N(0, 1), \quad (4.10)$$

where

$$\nu = c_1 \tau (p+q)(p+q+2) + c_2 \bar{\tau} q(q+2) \quad (4.11)$$

and

$$\begin{aligned} \sigma^2 = & 8\tau \left\{ c_1^2 (p+q)(p+q+2)(p+q+3) + c_1^2 p(p+q+2)^2 \right. \\ & \left. + q(c_1 \tau (p+q+2) + c_2 \bar{\tau} (q+2))^2 \right. \\ & \left. - 2c_1 (p+q+2) (pc_1 (p+q+2) + q(c_1 \tau (p+q+2) + c_2 \bar{\tau} (q+2))) \right\} \\ & + 8\bar{\tau} \left\{ c_2^2 q(q+2)(q+3) + q(c_1 \tau (p+q+2) + c_2 \bar{\tau} (q+2))^2 \right. \\ & \left. - 2c_2 (q+2) q(c_1 \tau (p+q+2) + c_2 \bar{\tau} (q+2)) \right\}. \end{aligned} \quad (4.12)$$

Remark 4.3. For $(c_1, c_2) = (\tau, \bar{\tau})$, (4.11) and (4.12) reduce to

$$\nu = \tau^2 (p+q)(p+q+2) + \bar{\tau}^2 q(q+2) \quad (4.13)$$

and

$$\sigma^2 = 8\{\tau^3(p+q)(p+q+2) + \bar{\tau}^3q(q+2) + \tau\bar{\tau}(\tau(p+q+2) - \bar{\tau}(q+2))^2\}, \quad (4.14)$$

respectively. For $(c_1, c_2) = (1, 1)$, (4.11) and (4.12) reduce, respectively, to

$$\nu = \tau(p+q)(p+q+2) + \bar{\tau}q(q+2) \quad (4.15)$$

and

$$\sigma^2 = 8\{\tau(p+q)(p+q+2) + \bar{\tau}q(q+2) + \tau\bar{\tau}p^2q\}. \quad (4.16)$$

Corollary 4.2 also generalizes the result of Mardia (1970) for complete data; indeed, when we set $\tau = 1$, i.e., $n = N$ in (4.15) and (4.16), we find that ν reduces to $(p+q)(p+q+2)$ and σ^2 reduces to $8(p+q)(p+q+2)$; hence (4.10) reduces to the result of Mardia.

4.1 Application to the Pennsylvania cholesterol data set

Assume that the Pennsylvania cholesterol data (Ryan, et al. 2005, p. 267) consist of mutually independent vectors and that missing observations are MCAR. For that data set, we have $p = 1$, $q = 2$, $N = 28$, and $n = 19$, and we choose $c_1 = \tau$ and $c_2 = \bar{\tau}$. The asymptotic mean and variance of $b_{2,1,2}$ are obtained from (4.11) and (4.12) to be $\nu = 7.7334$ and $\sigma^2 = 181.1658$, respectively. By Corollary 4.2, the asymptotic null distribution of the statistic (4.10) is

$$\sqrt{28}(b_{2,1,2} - 7.7334)/13.4598 \approx N(0, 1).$$

We calculate using (3.8) that the observed value of $b_{2,1,2}$ is 5.8623. Therefore, the observed value of the statistic (4.10) is

$$\sqrt{28}(5.8623 - 7.7334)/13.4598 = -0.7356.$$

The approximate P -value of the test is $2\Phi(-0.7356) = 0.2310$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. Therefore, we fail to reject the null hypothesis of multivariate normality at the 5% level of significance.

We note that the same conclusion is obtained by applying the classical Mardia statistic to the subset of the Pennsylvania cholesterol data set consisting of the $n = 19$ complete observations only. For this subset, the observed value of Mardia's statistic is 7.8176, the corresponding observed value of the normal approximation to Mardia's statistic is -0.1207 , hence the resulting approximate P -value for Mardia test is $2\Phi(-0.1207) = 0.9038$. However, this P -value is so large that the test based on the complete data appears unable to assess the strength of the evidence against the null hypothesis of normality, and this reflects the loss of information inherent in discarding the incomplete observations; cf., Little and Rubin (2002), p. 41. By contrast, the P -value based on the full data set appears to provide some measure of the strength of the evidence against the null hypothesis, even though the strength of that evidence is assessed to be too weak to reject that hypothesis.

We remark also that, in the case of the cholesterol data, the smaller sample size is less likely to yield an accurate normal approximation to Mardia's statistic, so the substantially larger P -value of the Mardia test for the complete data perhaps should be applied cautiously.

Acknowledgments. We are grateful to the referees and the associate editor for comments which helped us to improve the manuscript. The research of Richards was also supported by a 2013–2014

sabbatical leave-of-absence and a Romberg Guest Professorship at the Heidelberg University Graduate School for Mathematical and Computational Methods in the Sciences, funded by the German Universities Excellence Initiative grant GSC 220/2.

References

- Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis* (third edition). Wiley, New York.
- Chang, W.-Y., and Richards, D. St. P. (2009). Finite-sample inference with monotone incomplete multivariate normal data, I. *J. Multivariate Anal.*, **100**, 1883–1899.
- Chang, W.-Y., and Richards, D. St. P. (2010). Finite-sample inference with monotone incomplete multivariate normal data, II. *J. Multivariate Anal.*, **101**, 603–620.
- Cohn, N., Davidov, O., and Haitovsky, Y. (2008). Double sampling designs in multivariate linear models with missing data. *Comm. Statist. Simulation Comput.*, **37**, 1156–1166.
- Davidov, O., and Peddada, S. D. (2013). The linear stochastic order and directed inference for multivariate ordered distributions. *Ann. Statist.*, **41**, 1–40.
- Eaton, M. L., and Kariya, T. (1983). Multivariate tests with incomplete data. *Ann. Statist.*, **11**, 654–665.
- Garren, S. T., and Peddada, S. D. (2000). Asymptotic normality in multivariate nonlinear regression and multivariate generalized linear regression models under repeated measurements with missing data. *Statist. Probab. Lett.*, **48**, 293–302.
- Hao, J. and Krishnamoorthy, K. (2001). Inferences on a normal covariance matrix and generalized variance with monotone missing data. *J. Multivariate Anal.*, **78**, 62–82.
- Henze, N. (1994). On Mardia’s kurtosis test for multivariate normality. *Commun. Statist. – Theory & Methods*, **23**, 1031–1045.
- Henze, N. (2002). Invariant tests for multivariate normality: A critical review. *Statist. Papers*, **43**, 467–506.
- Krishnamoorthy, K., and Yu, J. (2012). Multivariate Behrens-Fisher problem with missing data. *J. Multivariate Anal.*, **105**, 141–150.
- Little, R. J. A. (1976). Inference about means from incomplete multivariate data, *Biometrika* 63 (1976) 593–604.
- Little, R. J. A., and Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (second edition). Wiley, Hoboken, NJ.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**, 519–530.
- Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhyā B*, **36**, 115–128.
- Peddada, S. D., Harris, S., Davidov, O. (2010). Analysis of correlated gene expression data on ordered categories. *J. Ind. Soc. Agric. Statist.*, **64**, 45–60.
- Richards, D. St. P., and Yamada, T. (2010). The Stein phenomenon for monotone incomplete multivariate normal data. *J. Multivariate Anal.*, **101**, 657–678.
- Romer, M. M. (2009). *The Statistical Analysis of Monotone Incomplete Multivariate Normal Data*. Doctoral Dissertation, Penn State University.

Romer, M. M., and Richards, D. St. P. (2010). Maximum likelihood estimation of the mean of a multivariate normal population with monotone incomplete data. *Statist. & Probab. Lett.*, **80**, 1284–1288.

Romer, M. M., and Richards, D. St. P. (2013). Finite-sample inference with monotone incomplete multivariate normal data, III: Hotelling's T^2 -statistic. *Statist. Modelling*, **13**, 431–457.

Ryan, B., Joiner, B., and Cryer, J. (2005). *Minitab Handbook* (fifth edition). Duxbury Press, Boston.

Yamada, T. (2013). Asymptotic properties of canonical correlation analysis for one group with additional observations. *J. Multivariate Anal.*, **114**, 389–401.

A Asymptotic expansions of $\widehat{\Sigma}$ and $\widehat{\Sigma}^{-1}$

To derive the expansions, we utilize results of Henze (1994). For a sequence of random vectors $\boldsymbol{\omega}_n$ and positive scalars a_n , we use the notation $\boldsymbol{\omega}_n = O_p(a_n)$ to denote that the sequence $\boldsymbol{\omega}_n/a_n$ is *tight*, and we write $\boldsymbol{\omega}_n = o(a_n)$ to mean that $\boldsymbol{\omega}_n/a_n$ converges to zero in probability.

Let $\mathbf{S} = n^{-1}\mathbf{A}$, where \mathbf{A} is given in (2.4). Since \mathbf{A} is the matrix of squares and cross-products of the n complete observations in the data set (1.1) then, following Henze (1994, p. 1035), we deduce that $\mathbf{S} = \mathbf{I}_{p+q} + n^{-1/2}\mathbf{B} + O_p(n^{-1})$, where

$$\mathbf{B} = n^{1/2} \left(\frac{1}{n} \sum_{j=1}^n \mathbf{Z}_j \mathbf{Z}_j' - \mathbf{I}_{p+q} \right) = O_p(1).$$

Write $\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix}$, where

$$\mathbf{B}_{11} = n^{1/2} \left(\frac{1}{n} \sum_{j=1}^n \mathbf{X}_j \mathbf{X}_j' - \mathbf{I}_p \right),$$

$$\mathbf{B}_{12} = \mathbf{B}_{21}' = n^{-1/2} \sum_{j=1}^n \mathbf{X}_j \mathbf{Y}_j',$$

$$\mathbf{B}_{22} = n^{1/2} \left(\frac{1}{n} \sum_{j=1}^n \mathbf{Y}_j \mathbf{Y}_j' - \mathbf{I}_q \right),$$

and define $\mathbf{B}^* = \tau^{-1/2}\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \mathbf{B}_{22}^* \end{pmatrix}$. Then,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22,n} \end{pmatrix} = n\mathbf{I}_{p+q} + n^{1/2}\mathbf{B} + O_p(1),$$

so that

$$\begin{aligned} \mathbf{A}_{11} &= n\mathbf{I}_p + n^{1/2}\mathbf{B}_{11} + O_p(1), \\ \mathbf{A}_{12} &= \mathbf{A}_{21}' = n^{1/2}\mathbf{B}_{12} + O_p(1), \\ \mathbf{A}_{22,n} &= n\mathbf{I}_q + n^{1/2}\mathbf{B}_{22} + O_p(1). \end{aligned} \tag{A.1}$$

Next, we follow Henze (1994, p. 1035, equation (2.2)) to obtain

$$\mathbf{A}_{22,n}^{-1} = n^{-1}\mathbf{I}_q - n^{-3/2}\mathbf{B}_{22} + O_p(n^{-2}).$$

By the same argument, we obtain

$$\mathbf{A}_{22,N} = N\mathbf{I}_q + N^{1/2}\widetilde{\mathbf{B}}_{22} + O_p(1),$$

and, by inversion,

$$\mathbf{A}_{22,N}^{-1} = N^{-1}\mathbf{I}_q - N^{-3/2}\tilde{\mathbf{B}}_{22} + O_p(N^{-2}),$$

where

$$\tilde{\mathbf{B}}_{22} = N^{1/2}\left(\frac{1}{N}\sum_{j=1}^N \mathbf{Y}_j \mathbf{Y}_j' - \mathbf{I}_q\right) = O_p(1).$$

The *partial Iwasawa coordinates* of $\hat{\Sigma}$ (Chang and Richards (2009, equation (4.5)) are

$$\hat{\Delta}_{11} = \frac{1}{n}\mathbf{A}_{11 \cdot 2, n}, \quad \hat{\Delta}_{12} = \hat{\Delta}_{21}' = \mathbf{A}_{12}\mathbf{A}_{22, n}^{-1}, \quad \hat{\Delta}_{22} = \frac{1}{N}\mathbf{A}_{22, N}. \quad (\text{A.2})$$

These matrices are such that

$$\hat{\Sigma} = \begin{pmatrix} \mathbf{I}_p & \hat{\Delta}_{12} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \begin{pmatrix} \hat{\Delta}_{11} & \mathbf{0} \\ \mathbf{0} & \hat{\Delta}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \hat{\Delta}_{21} & \mathbf{I}_q \end{pmatrix}, \quad (\text{A.3})$$

hence

$$\hat{\Sigma}^{-1} = \begin{pmatrix} \hat{\Delta}_{11}^{-1} & -\hat{\Delta}_{11}^{-1}\hat{\Delta}_{12} \\ -\hat{\Delta}_{21}\hat{\Delta}_{11}^{-1} & \hat{\Delta}_{22}^{-1} + \hat{\Delta}_{21}\hat{\Delta}_{11}^{-1}\hat{\Delta}_{12} \end{pmatrix}. \quad (\text{A.4})$$

Noting that $n = \tau N$ and $\mathbf{B}^* = \tau^{-1/2}\mathbf{B}$, we apply (A.1) to deduce that $\hat{\Delta}_{11}$, $\hat{\Delta}_{12}$, and $\hat{\Delta}_{22}$ have the following asymptotic expansions:

$$\begin{aligned} \hat{\Delta}_{11} &= \frac{1}{n}\{\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22, n}^{-1}\mathbf{A}_{21}\} \\ &= \frac{1}{n}\{n\mathbf{I}_p + n^{1/2}\mathbf{B}_{11} - (n^{1/2}\mathbf{B}_{12})(n^{-1}\mathbf{I}_q - n^{-3/2}\mathbf{B}_{22})(n^{1/2}\mathbf{B}_{21})\} + O_p(n^{-1}) \\ &= \mathbf{I}_p + n^{-1/2}\mathbf{B}_{11} + O_p(n^{-1}) \\ &= \mathbf{I}_p + N^{-1/2}\mathbf{B}_{11}^* + O_p(N^{-1}); \end{aligned}$$

$$\begin{aligned} \hat{\Delta}_{12} &= \mathbf{A}_{12}\mathbf{A}_{22, n}^{-1} = n^{1/2}\mathbf{B}_{12}(n^{-1}\mathbf{I}_q - n^{-3/2}\mathbf{B}_{22}) + O_p(n^{-1}) \\ &= n^{-1/2}\mathbf{B}_{12} - n^{-1}\mathbf{B}_{12}\mathbf{B}_{22} + O_p(n^{-1}) \\ &= N^{-1/2}\mathbf{B}_{12}^* - N^{-1}\mathbf{B}_{12}^*\mathbf{B}_{22}^* + O_p(N^{-1}); \end{aligned}$$

and

$$\hat{\Delta}_{22} = \frac{1}{N}\mathbf{A}_{22, N} = \mathbf{I}_q + N^{-1/2}\tilde{\mathbf{B}}_{22} + O_p(N^{-1}).$$

By the same inversion argument used earlier, we also have

$$\hat{\Delta}_{11}^{-1} = \mathbf{I}_p - N^{-1/2}\mathbf{B}_{11}^* + O_p(N^{-1}), \quad \hat{\Delta}_{22}^{-1} = \mathbf{I}_q - N^{-1/2}\tilde{\mathbf{B}}_{22} + O_p(N^{-1}).$$

Collecting together these results, we obtain

$$\begin{aligned} \hat{\Sigma}_{11} &= \hat{\Delta}_{11} + \hat{\Delta}_{12}\hat{\Delta}_{22}^{-1}\hat{\Delta}_{21} = \mathbf{I}_p + N^{-1/2}\mathbf{B}_{11}^* + O_p(N^{-1}), \\ \hat{\Sigma}_{12} &= \hat{\Delta}_{12}\hat{\Delta}_{22} = N^{-1/2}\mathbf{B}_{12}^* + O_p(N^{-1}), \\ \hat{\Sigma}_{22} &= \hat{\Delta}_{22} = \mathbf{I}_q + N^{-1/2}\tilde{\mathbf{B}}_{22} + O_p(N^{-1}). \end{aligned}$$

Therefore,

$$\widehat{\Sigma} = \mathbf{I}_{p+q} + N^{-1/2} \widetilde{\mathbf{B}} + O_p(N^{-1}), \quad (\text{A.5})$$

where $\widetilde{\mathbf{B}} = \begin{pmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \widetilde{\mathbf{B}}_{22} \end{pmatrix}$. By inverting (A.5), we obtain the asymptotic expansion,

$$\widehat{\Sigma}^{-1} = \mathbf{I}_{p+q} - N^{-1/2} \widetilde{\mathbf{B}} + O_p(N^{-1}). \quad (\text{A.6})$$

B Asymptotic expansions of $b_{2,p,q}^{(1)}$ and $b_{2,p,q}^{(2)}$

The statistic $b_{2,p,q}^{(2)}$ is reminiscent of a statistic studied by Henze (1994) so we shall proceed similarly to Henze in obtaining an asymptotic expansion for $b_{2,p,q}^{(2)}$. Nevertheless, we note that the statistic $b_{2,p,q}^{(2)}$ is not identical with Henze's statistic; indeed, although $b_{2,p,q}^{(2)}$ is given by a sum over the indices $j = n+1, \dots, N$, we note that $\bar{\mathbf{Y}}$ and $\widehat{\Sigma}_{22} = N^{-1} \mathbf{A}_{22,N}$ are the sample mean and sample covariance matrix, respectively, of $\mathbf{Y}_1, \dots, \mathbf{Y}_N$, so that $b_{2,p,q}^{(2)}$ incorporates information from the portion of the data that is complete.

Thus, by direct expansion, we have

$$\begin{aligned} [(\mathbf{Y}_j - \bar{\mathbf{Y}})' \mathbf{A}_{22,N}^{-1} (\mathbf{Y}_j - \bar{\mathbf{Y}})]^2 &= (\mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j - 2\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j + \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}})^2 \\ &= (\mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j)^2 + 4(\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j)^2 \\ &\quad + (\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}})^2 - 4\mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \\ &\quad + 2\mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}} - 4\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}}. \end{aligned} \quad (\text{B.1})$$

With Ξ^* and Θ^* defined as in (4.3) we obtain, by applying the methods of Henze (1994), the approximations

$$\frac{1}{N} \sum_{j=n+1}^N (\mathbf{Y}_j \mathbf{Y}_j')^2 = \bar{\tau} \Xi^* + o(1)$$

and

$$\frac{1}{N} \sum_{j=n+1}^N \|\mathbf{Y}_j\|^2 \mathbf{Y}_j = \bar{\tau} \Theta^* + o(1).$$

Therefore,

$$\frac{1}{N} \sum_{j=n+1}^N (\mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j)^2 = \frac{1}{N} \sum_{j=n+1}^N (\mathbf{Y}_j' \mathbf{Y}_j)^2 - 2N^{-1/2} \bar{\tau} \text{tr}(\widetilde{\mathbf{B}}_{22} \Xi^*) + o(N^{-1/2}),$$

and

$$\frac{1}{N} \sum_{j=n+1}^N \mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j = \bar{\tau} \Theta^{*'} \bar{\mathbf{Y}} + o(N^{-1/2}).$$

Also,

$$N^{-1} \sum_{j=n+1}^N (\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j)^2 = O_p(N^{-1}),$$

$$N^{-1} \sum_{j=n+1}^N (\bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}})^2 = O_p(N^{-2}),$$

$$N^{-1} \sum_{j=n+1}^N \mathbf{Y}_j' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}} = O_p(N^{-1}),$$

and

$$N^{-1} \sum_{j=n+1}^N \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \mathbf{Y}_j \bar{\mathbf{Y}}' \mathbf{A}_{22,N}^{-1} \bar{\mathbf{Y}} = O_p(N^{-1}).$$

Summing (B.1) over $j = n+1, \dots, N$, we obtain

$$\frac{1}{N} b_{2,p,q}^{(2)} = \frac{1}{N} \sum_{j=n+1}^N (\mathbf{Y}_j' \mathbf{Y}_j)^2 - 2N^{-1/2} \bar{\tau} \text{tr}(\tilde{\mathbf{B}}_{22} \boldsymbol{\Xi}^*) - 4\bar{\tau} \boldsymbol{\Theta}^{*'} \bar{\mathbf{Y}} + o(N^{-1/2}). \quad (\text{B.2})$$

Consider next the statistic $b_{2,p,q}^{(1)}$. We apply (3.7) to write

$$\begin{aligned} [(\mathbf{Z}_j - \hat{\boldsymbol{\mu}})' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \hat{\boldsymbol{\mu}})]^2 &= [(\mathbf{Z}_j - \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\mathbf{Y}})' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\mathbf{Y}})]^2 \\ &= [(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) + 2\bar{\tau} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}} + \bar{\tau}^2 \tilde{\mathbf{Y}}' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}}]^2, \end{aligned}$$

expand this expression directly, and sum over $j = 1, \dots, n$. By (3.5), $\sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) = \mathbf{0}$, so we obtain $b_{2,p,q}^{(1)}$ as a sum of five terms, $b_{2,p,q}^{(1)} = \sum_{j=1}^5 T_j$, where

$$\begin{aligned} T_1 &= \sum_{j=1}^n [(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)]^2, \\ T_2 &= 4\bar{\tau} \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}}, \\ T_3 &= 4\bar{\tau}^2 \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}} \tilde{\mathbf{Y}}' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1), \\ T_4 &= \bar{\tau}^2 \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) \tilde{\mathbf{Y}}' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}}, \\ T_5 &= n\bar{\tau}^2 (\tilde{\mathbf{Y}}' \hat{\boldsymbol{\Sigma}}^{-1} \tilde{\mathbf{Y}})^2. \end{aligned}$$

From (2.4),

$$\sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' = \mathbf{A} \equiv \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22,n} \end{pmatrix}$$

and by (A.4),

$$\hat{\boldsymbol{\Sigma}}^{-1} = \begin{pmatrix} n\mathbf{A}_{11,2}^{-1} & -n\mathbf{A}_{11,2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22,n}^{-1} \\ -n\mathbf{A}_{22,n}^{-1} \mathbf{A}_{21} \mathbf{A}_{11,2}^{-1} & N\mathbf{A}_{22,N}^{-1} + n\mathbf{A}_{22,n}^{-1} \mathbf{A}_{21} \mathbf{A}_{11,2}^{-1} \mathbf{A}_{12} \mathbf{A}_{22,n}^{-1} \end{pmatrix}.$$

By direct multiplication, we obtain

$$\hat{\boldsymbol{\Sigma}}^{-1} \mathbf{A} = \begin{pmatrix} n\mathbf{I}_p & \mathbf{0} \\ n\mathbf{A}_{22,n}^{-1} \mathbf{A}_{21} + N\mathbf{A}_{22,N}^{-1} \mathbf{A}_{21} & N\mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \end{pmatrix} \quad (\text{B.3})$$

and

$$\mathbf{A} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{A} = \begin{pmatrix} n\mathbf{A}_{11,2,n} + N\mathbf{A}_{12} \mathbf{A}_{22,N}^{-1} \mathbf{A}_{21} & N\mathbf{A}_{12} \mathbf{A}_{22,n}^{-1} \mathbf{A}_{22,n} \\ N\mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} \mathbf{A}_{21} & N\mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \end{pmatrix}, \quad (\text{B.4})$$

and then it follows from (3.6) that

$$\begin{aligned}\tilde{\mathbf{Y}}\hat{\Sigma}^{-1}\tilde{\mathbf{Y}} &= \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \mathbf{A}\hat{\Sigma}^{-1}\mathbf{A} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix} \\ &= N(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2).\end{aligned}$$

Also, $N^{1/2}(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' = O_p(N^{-1/2})$.

The statistic T_1 is the same as in the complete case, so we apply the expansion derived by Henze (1994). Letting $\Xi = E((\mathbf{Z}\mathbf{Z}')^2)$ and $\Theta = E(\|\mathbf{Z}\|^2 \mathbf{Z})$, we obtain

$$\begin{aligned}\frac{1}{N}T_1 &= \frac{1}{N} \sum_{j=1}^n [(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)]^2 \\ &= \frac{1}{N} \sum_{j=1}^n (\mathbf{Z}_j' \mathbf{Z}_j)^2 - 2N^{-1/2} \tau \text{tr}(\tilde{\mathbf{B}}\Xi) - 4\tau \Theta' \bar{\mathbf{Z}}_1 + o(N^{-1/2}).\end{aligned}$$

We now consider T_2 . Letting $\tilde{\mathbf{Y}}^* = \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}$, we obtain

$$\tilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{0} & \mathbf{A}_{12}\mathbf{A}_{22,n}^{-1} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \tilde{\mathbf{Y}}^* = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \tilde{\mathbf{Y}}^* + N^{-1/2} \begin{pmatrix} \mathbf{0} & \mathbf{B}_{12} \\ \mathbf{0} & \mathbf{I}_q \end{pmatrix} \tilde{\mathbf{Y}}^* + O_p(N^{-1}),$$

and

$$\begin{aligned}(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\ = \mathbf{Z}_j' \hat{\Sigma}^{-1} \mathbf{Z}_j \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} - \mathbf{Z}_j' \hat{\Sigma}^{-1} \mathbf{Z}_j \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} - 2\mathbf{Z}_j' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\ + 2\mathbf{Z}_j' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} + \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} - \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}}.\end{aligned}$$

Then,

$$\begin{aligned}\frac{1}{N} \sum_{j=1}^n \mathbf{Z}_j' \hat{\Sigma}^{-1} \mathbf{Z}_j \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} &= \tau \Theta' \tilde{\mathbf{Y}}^* + o(N^{-1/2}), \\ \frac{1}{N} \sum_{j=1}^n \mathbf{Z}_j' \hat{\Sigma}^{-1} \mathbf{Z}_j \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} &= O_p(N^{-1}), \\ \frac{1}{N} \sum_{j=1}^n \mathbf{Z}_j' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} &= O_p(N^{-1}),\end{aligned}$$

and

$$\begin{aligned}\frac{1}{N} \sum_{j=1}^n \{2\mathbf{Z}_j' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} + \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \mathbf{Z}_j' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} - \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}}\} \\ = 2\tau \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \bar{\mathbf{Z}} \bar{\mathbf{Z}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} = O_p(N^{-2}).\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{1}{4\bar{\tau}N}T_2 &= \frac{1}{N} \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\ &= \tau \Theta' \tilde{\mathbf{Y}}^* + o(N^{-1/2}).\end{aligned}$$

Next,

$$\begin{aligned}
\frac{1}{4\bar{\tau}^2}T_3 &= \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) \\
&= \tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} \left[\sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \right] \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\
&= \tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} \mathbf{A} \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\
&= \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}' \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \mathbf{A} \hat{\Sigma}^{-1} \mathbf{A} \hat{\Sigma}^{-1} \mathbf{A} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}.
\end{aligned}$$

It may be shown using (B.3) and (B.4) that

$$\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \mathbf{A} \hat{\Sigma}^{-1} \mathbf{A} \hat{\Sigma}^{-1} \mathbf{A} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} = \begin{pmatrix} * & * \\ * & N^2 \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} \end{pmatrix},$$

where “*” denotes terms whose explicit expressions are not needed. Consequently,

$$\begin{aligned}
\frac{1}{4\bar{\tau}^2}T_3 &= \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}' \begin{pmatrix} * & * \\ * & N^2 \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22,n}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix} \\
&= N^2 (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) \\
&= O_p(1).
\end{aligned}$$

As for T_4 , we have

$$\begin{aligned}
\frac{1}{\bar{\tau}^2}T_4 &= \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) \tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\
&= \tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \cdot \text{tr} \hat{\Sigma}^{-1} \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \\
&= N (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) \text{tr} \hat{\Sigma}^{-1} \mathbf{A}.
\end{aligned}$$

Applying (B.3), we obtain

$$\frac{1}{\bar{\tau}^2}T_4 = N (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) (n \text{tr} \mathbf{I}_p + N \text{tr} \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n}) = O_p(1).$$

In the case of T_5 , we find that

$$T_5 = n \bar{\tau}^2 (\tilde{\mathbf{Y}}' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}})^2 = n N \bar{\tau}^2 [(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)]^2 = O_p(N^{-2})$$

because $\mathbf{A}_{22,N}^{-1} = O_p(N^{-1})$ and $(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' = O_p(N^{-1})$.

Collecting together these five expansions, we deduce that

$$\begin{aligned}
\frac{1}{N}b_{2,p,q}^{(1)} &= N^{-1} \sum_{j=1}^5 T_j \\
&= N^{-1} \left\{ \sum_{j=1}^n [(\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)]^2 \right. \\
&\quad + 4\bar{\tau} \sum_{j=1}^n (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} (\mathbf{Z}_j - \bar{\mathbf{Z}}_1) (\mathbf{Z}_j - \bar{\mathbf{Z}}_1)' \hat{\Sigma}^{-1} \tilde{\mathbf{Y}} \\
&\quad + 4\bar{\tau}^2 N^2 (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n} \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) \\
&\quad + \bar{\tau}^2 N (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) (np + N \text{tr} \mathbf{A}_{22,N}^{-1} \mathbf{A}_{22,n}) \\
&\quad \left. + \bar{\tau}^3 n N^2 [(\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)' \mathbf{A}_{22,N}^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)]^2 \right\} \\
&= N^{-1} \sum_{j=1}^n (\mathbf{Z}'_j \mathbf{Z}_j)^2 - 2N^{-1/2} \tau \text{tr}(\tilde{\mathbf{B}} \Xi) - 4\tau \Theta' \bar{\mathbf{Z}} + 4\tau \bar{\tau} \Theta' \tilde{\mathbf{Y}}^* + o(N^{-1/2}). \tag{B.5}
\end{aligned}$$

Finally, it follows from (B.2) and (B.5) that

$$\begin{aligned}
b_{2,p,q} &= N^{-1} \left(c_1 b_{2,p,q}^{(1)} + c_2 b_{2,p,q}^{(2)} \right) \\
&= c_1 \left\{ N^{-1} \sum_{j=1}^n (\mathbf{Z}'_j \mathbf{Z}_j)^2 - 2N^{-1/2} \tau \text{tr}(\tilde{\mathbf{B}} \Xi) - 4\tau \Theta' \bar{\mathbf{Z}}_1 + 4\tau \bar{\tau} \Theta' \tilde{\mathbf{Y}}^* \right\} \\
&\quad + c_2 \left\{ N^{-1} \sum_{j=n+1}^N (\mathbf{Y}'_j \mathbf{Y}_j)^2 - 2N^{-1/2} \bar{\tau} \text{tr}(\tilde{\mathbf{B}}_{22} \Xi^*) - 4\bar{\tau} \Theta^{*'} \tilde{\mathbf{Y}} \right\} + o(N^{-1/2}). \tag{B.6}
\end{aligned}$$

C The asymptotic distribution of $b_{2,p,q}$

C.1 The proof of Theorem 4.1

To obtain the asymptotic distribution of $b_{2,p,q}$, we first simplify the expression (B.6). Recall that

$$\mathbf{B}^* = \begin{pmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \mathbf{B}_{22}^* \end{pmatrix}$$

and

$$\tilde{\mathbf{B}} = \begin{pmatrix} \mathbf{B}_{11}^* & \mathbf{B}_{12}^* \\ \mathbf{B}_{21}^* & \tilde{\mathbf{B}}_{22} \end{pmatrix}$$

where

$$\begin{aligned}
\mathbf{B}_{11}^* &= N^{1/2} \left(\frac{1}{n} \sum_{j=1}^n \mathbf{X}_j \mathbf{X}'_j - \mathbf{I}_p \right), \quad \mathbf{B}_{12}^* = \mathbf{B}_{21}^* = N^{1/2} \frac{1}{n} \sum_{j=1}^n \mathbf{X}_j \mathbf{Y}'_j, \\
\mathbf{B}_{22}^* &= N^{1/2} \left(\frac{1}{n} \sum_{j=1}^n \mathbf{Y}_j \mathbf{Y}'_j - \mathbf{I}_q \right), \quad \tilde{\mathbf{B}}_{22} = N^{1/2} \left(\frac{1}{N} \sum_{j=1}^N \mathbf{Y}_j \mathbf{Y}'_j - \mathbf{I}_q \right).
\end{aligned}$$

Define

$$\tilde{\mathbf{B}}_{22}^* = \sqrt{N} \left(\frac{1}{N-n} \sum_{j=n+1}^N \mathbf{Y}_j \mathbf{Y}'_j - \mathbf{I}_q \right),$$

then $\tilde{\mathbf{B}}_{22} = \tau \mathbf{B}_{22}^* + \bar{\tau} \tilde{\mathbf{B}}_{22}^*$. Recalling from (4.1) and (4.2) the definitions of the matrix Ξ and its components Ξ_{ij} , direct algebraic calculations reveal that

$$\begin{aligned} & c_1 \tau \operatorname{tr}(\tilde{\mathbf{B}} \Xi) + c_2 \bar{\tau} \operatorname{tr}(\tilde{\mathbf{B}}_{22} \Xi^*) \\ &= c_1 \tau \{ \operatorname{tr}(\mathbf{B}_{11}^* \Xi_{11} + \mathbf{B}_{12}^* \Xi_{21}) + \operatorname{tr}(\mathbf{B}_{21}^* \Xi_{12} + (\tau \mathbf{B}_{22}^* + \bar{\tau} \tilde{\mathbf{B}}_{22}^*) \Xi_{22}) \} \\ & \quad + c_2 \bar{\tau} \operatorname{tr}((\tau \mathbf{B}_{22}^* + \bar{\tau} \tilde{\mathbf{B}}_{22}^*) \Xi^*) \\ &= c_1 \tau \{ \operatorname{tr}(\mathbf{B}_{11}^* \Xi_{11} + \mathbf{B}_{12}^* \Xi_{21}) + \operatorname{tr}(\mathbf{B}_{21}^* \Xi_{12} + \tau \mathbf{B}_{22}^* \Xi_{22}) \} + c_2 \bar{\tau} \operatorname{tr}(\tau \mathbf{B}_{22}^* \Xi^*) \\ & \quad + c_1 \tau \operatorname{tr}(\bar{\tau} \tilde{\mathbf{B}}_{22}^* \Xi_{22}) + c_2 \bar{\tau} \operatorname{tr}(\bar{\tau} \tilde{\mathbf{B}}_{22}^* \Xi^*) \\ &\equiv \tau \operatorname{tr} \mathbf{B}^* \tilde{\Xi} + \bar{\tau} \operatorname{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\Xi}_{22}, \end{aligned}$$

Since $\bar{\mathbf{Z}}_1 = \begin{pmatrix} \bar{\mathbf{X}} \\ \bar{\mathbf{Y}}_1 \end{pmatrix}$, $\tilde{\mathbf{Y}}^* = \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2 \end{pmatrix}$, and $\bar{\mathbf{Y}} = \tau \bar{\mathbf{Y}}_1 + \bar{\tau} \bar{\mathbf{Y}}_2$, then it follows from the definitions of Θ and $\tilde{\Theta}$ in (4.5) and (4.6), respectively, that

$$\begin{aligned} & c_1 \tau \Theta' \bar{\mathbf{Z}}_1 - c_1 \tau \bar{\tau} \Theta' \tilde{\mathbf{Y}}^* + c_2 \bar{\tau} \Theta'^* \bar{\mathbf{Y}} \\ &= c_1 \tau (\Theta'_1 \bar{\mathbf{X}} + \Theta'_2 \bar{\mathbf{Y}}_1) - c_1 \tau \bar{\tau} \Theta'_2 (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2) + c_2 \bar{\tau} \Theta'^* (\tau \bar{\mathbf{Y}}_1 + \bar{\tau} \bar{\mathbf{Y}}_2) \\ &= c_1 \tau \Theta'_1 \bar{\mathbf{X}} + (c_1 \tau \Theta_2 - c_1 \tau \bar{\tau} \Theta_2 + c_2 \tau \bar{\tau} \Theta^*)' \bar{\mathbf{Y}}_1 + (c_1 \tau \bar{\tau} \Theta_2 + c_2 \bar{\tau}^2 \Theta^*)' \bar{\mathbf{Y}}_2 \\ &= c_1 \tau \Theta'_1 \bar{\mathbf{X}} + \tau (c_1 \tau \Theta_2 + c_2 \bar{\tau} \Theta^*)' \bar{\mathbf{Y}}_1 + \bar{\tau} (c_1 \tau \Theta_2 + c_2 \bar{\tau} \Theta^*)' \bar{\mathbf{Y}}_2 \\ &\equiv \tau \tilde{\Theta}' \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\Theta}'_2 \bar{\mathbf{Y}}_2. \end{aligned}$$

Hence,

$$\begin{aligned} b_{2,p,q} &= \frac{c_1}{N} \sum_{j=1}^n (\mathbf{Z}'_j \mathbf{Z}_j)^2 + \frac{c_2}{N} \sum_{j=n+1}^N (\mathbf{Y}'_j \mathbf{Y}_j)^2 \\ &\quad - 2N^{-1/2} (\tau \operatorname{tr} \mathbf{B}^* \tilde{\Xi} + \bar{\tau} \operatorname{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\Xi}_{22}) - 4(\tau \tilde{\Theta}' \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\Theta}'_2 \bar{\mathbf{Y}}_2) + o(N^{-1/2}). \end{aligned}$$

Since $\tau = n/N$ and $\bar{\tau} = (N-n)/N$ then $N^{-1/2} = n^{-1/2} \tau^{1/2}$ and $N^{-1/2} = (N-n)^{-1/2} \bar{\tau}^{1/2}$. Define

$$\begin{pmatrix} U_1 \\ V_1 \\ \mathbf{W}_1 \end{pmatrix} = n^{-1/2} \sum_{j=1}^n \tilde{\mathbf{Z}}_j, \quad \tilde{\mathbf{Z}}_j = \begin{pmatrix} \|\mathbf{Z}_j\|^4 - E\|\mathbf{Z}\|^4 \\ \mathbf{Z}'_j \tilde{\Xi} \mathbf{Z}_j - E(\mathbf{Z}' \tilde{\Xi} \mathbf{Z}) \\ \mathbf{X}_j \\ \mathbf{Y}_j \end{pmatrix}, \quad (\text{C.1})$$

$$\begin{pmatrix} U_2 \\ V_2 \\ \mathbf{W}_2 \end{pmatrix} = (N-n)^{-1/2} \sum_{j=n+1}^N \tilde{\mathbf{Z}}_j^*, \quad \tilde{\mathbf{Z}}_j^* = \begin{pmatrix} \|\mathbf{Y}_j\|^4 - E\|\mathbf{Y}\|^4 \\ \mathbf{Y}'_j \tilde{\Xi} \mathbf{Y}_j - E(\mathbf{Y}' \tilde{\Xi}_{22} \mathbf{Y}) \\ \mathbf{Y}_j \end{pmatrix};$$

then,

$$\begin{aligned} & N^{1/2} (b_{2,p,q} - c_1 \tau E|\mathbf{Z}|^4 - c_2 \bar{\tau} E|\mathbf{Y}|^4) \\ &= N^{-1/2} \left\{ c_1 \sum_{j=1}^n (\mathbf{Z}'_j \mathbf{Z}_j)^2 + c_2 \sum_{j=n+1}^N (\mathbf{Y}'_j \mathbf{Y}_j)^2 \right\} - N^{1/2} \left\{ c_1 \tau E|\mathbf{Z}|^4 - c_2 \bar{\tau} E|\mathbf{Y}|^4 \right\} \\ & \quad - 2N^{-1/2} (\tau \operatorname{tr} \mathbf{B}^* \tilde{\Xi} + \bar{\tau} \operatorname{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\Xi}_{22}) - 4(\tau \tilde{\Theta}' \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\Theta}'_2 \bar{\mathbf{Y}}_2) + o(N^{-1/2}) \\ &= \tau^{1/2} n^{-1/2} c_1 \sum_{j=1}^n (\mathbf{Z}'_j \mathbf{Z}_j)^2 + \bar{\tau}^{1/2} (N-n)^{-1/2} c_2 \sum_{j=n+1}^N (\mathbf{Y}'_j \mathbf{Y}_j)^2 \\ & \quad - n^{1/2} c_1 \tau^{1/2} E|\mathbf{Z}|^4 - (N-n)^{1/2} c_2 \bar{\tau}^{1/2} E|\mathbf{Y}|^4 \\ & \quad - 2N^{-1/2} (\tau \operatorname{tr} \mathbf{B}^* \tilde{\Xi} + \bar{\tau} \operatorname{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\Xi}_{22}) - 4(\tau \tilde{\Theta}' \bar{\mathbf{Z}}_1 + \bar{\tau} \tilde{\Theta}'_2 \bar{\mathbf{Y}}_2) + o(N^{-1/2}). \end{aligned}$$

Writing each $(\mathbf{Z}'_j \mathbf{Z}_j)^2 = (\mathbf{Z}'_j \mathbf{Z}_j)^2 - E|\mathbf{Z}|^4 + E|\mathbf{Z}|^4$, and similarly for each $|\mathbf{Y}|^4$, we obtain

$$\begin{aligned}
& N^{1/2} (b_{2,p,q} - c_1 \tau E|\mathbf{Z}|^4 - c_2 \bar{\tau} E|\mathbf{Y}|^4) \\
&= \tau^{1/2} c_1 n^{-1/2} \sum_{j=1}^n ((\mathbf{Z}'_j \mathbf{Z}_j)^2 - E|\mathbf{Z}|^4) + \bar{\tau}^{1/2} (N-n)^{-1/2} c_2 \sum_{j=n+1}^N ((\mathbf{Y}'_j \mathbf{Y}_j)^2 - E|\mathbf{Y}|^4) \\
&\quad - 2N^{-1/2} (\tau \text{tr} \mathbf{B}^* \tilde{\mathbf{\Xi}} + \bar{\tau} \text{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\mathbf{\Xi}}_{22}) - 4(\tau \tilde{\mathbf{\Theta}}' \tilde{\mathbf{Z}}_1 + \bar{\tau} \tilde{\mathbf{\Theta}}'_2 \tilde{\mathbf{Y}}_2) + o(N^{-1/2}) \\
&= \tau^{1/2} c_1 U_1 + \bar{\tau}^{1/2} c_2 U_2 \\
&\quad - 2N^{-1/2} (\tau \text{tr} \mathbf{B}^* \tilde{\mathbf{\Xi}} + \bar{\tau} \text{tr} \tilde{\mathbf{B}}_{22}^* \tilde{\mathbf{\Xi}}_{22}) - 4(\tau \tilde{\mathbf{\Theta}}' \tilde{\mathbf{Z}}_1 + \bar{\tau} \tilde{\mathbf{\Theta}}'_2 \tilde{\mathbf{Y}}_2) + o(N^{-1/2}),
\end{aligned}$$

so we obtain

$$\begin{aligned}
N^{1/2} (b_{2,p,q} - c_1 \tau E\|\mathbf{Z}\|^4 - c_2 \bar{\tau} E\|\mathbf{Y}\|^4) &= c_1 \tau^{1/2} U_1 - 2\tau^{1/2} V_1 - 4\tau^{1/2} \tilde{\mathbf{\Theta}}' \mathbf{W}_1 \\
&\quad + c_2 \bar{\tau}^{1/2} U_2 - 2\bar{\tau}^{1/2} V_2 - 4\bar{\tau}^{1/2} \tilde{\mathbf{\Theta}}'_2 \mathbf{W}_2 + o(1) \\
&\equiv \tilde{b}_{2,p,q}^{(1)} + \tilde{b}_{2,p,q}^{(2)} + o(1),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{b}_{2,p,q}^{(1)} &= c_1 \tau^{1/2} U_1 - 2\tau^{1/2} V_1 - 4\tau^{1/2} \tilde{\mathbf{\Theta}}' \mathbf{W}_1, \\
\tilde{b}_{2,p,q}^{(2)} &= c_2 \bar{\tau}^{1/2} U_2 - 2\bar{\tau}^{1/2} V_2 - 4\bar{\tau}^{1/2} \tilde{\mathbf{\Theta}}'_2 \mathbf{W}_2.
\end{aligned} \tag{C.2}$$

Note that $\tilde{b}_{2,p,q}^{(1)}$ depends on $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ only, and $\tilde{b}_{2,p,q}^{(2)}$ depends on $\mathbf{Y}_{n+1}, \dots, \mathbf{Y}_N$ only. Since $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ and $\mathbf{Y}_{n+1}, \dots, \mathbf{Y}_N$ are independent, then $\tilde{b}_{2,p,q}^{(1)}$ and $\tilde{b}_{2,p,q}^{(2)}$ also are independent.

By (C.1), $\tilde{\mathbf{Z}}_1, \dots, \tilde{\mathbf{Z}}_n$ are mutually independent and identically distributed with $E(\tilde{\mathbf{Z}}_1) = \mathbf{0}$ and covariance matrix

$$E(\tilde{\mathbf{Z}}_1 \tilde{\mathbf{Z}}_1') = \begin{pmatrix} \text{Var}(\|\mathbf{Z}\|^4) & \text{Cov}(\|\mathbf{Z}\|^4, \mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) & E(\|\mathbf{Z}\|^4 \mathbf{Z}') \\ \text{Cov}(\|\mathbf{Z}\|^4, \mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) & \text{Var}(\mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) & E(\mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z} \mathbf{Z}') \\ E(\|\mathbf{Z}\|^4 \mathbf{Z}) & E(\mathbf{Z} \mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) & \mathbf{I}_{p+q} \end{pmatrix}, \tag{C.3}$$

and $\tilde{\mathbf{Z}}_{n+1}^*, \dots, \tilde{\mathbf{Z}}_N^*$ are mutually independent and identically distributed with $E(\tilde{\mathbf{Z}}_N^*) = \mathbf{0}$ and covariance matrix

$$E(\tilde{\mathbf{Z}}_N^* \tilde{\mathbf{Z}}_N^{*'}) = \begin{pmatrix} \text{Var}(\|\mathbf{Y}\|^4) & \text{Cov}(\|\mathbf{Y}\|^4, \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) & E(\|\mathbf{Y}\|^4 \mathbf{Y}') \\ \text{Cov}(\|\mathbf{Y}\|^4, \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) & \text{Var}(\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) & E(\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y} \mathbf{Y}') \\ E(\|\mathbf{Y}\|^4 \mathbf{Y}) & E(\mathbf{Y} \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) & \mathbf{I}_{p+q} \end{pmatrix}. \tag{C.4}$$

Since $\tilde{b}_{2,p,q}^{(1)}$ and $\tilde{b}_{2,p,q}^{(2)}$ are independent, it follows that

$$\begin{aligned}
\text{Var}(b_{2,p,q}) &= \text{Var}(\tilde{b}_{2,p,q}^{(1)}) + \text{Var}(\tilde{b}_{2,p,q}^{(2)}) \\
&= \tau \boldsymbol{\nu}'_1 E(\tilde{\mathbf{Z}}_1 \tilde{\mathbf{Z}}_1') \boldsymbol{\nu}_1 + \bar{\tau} \boldsymbol{\nu}'_2 E(\tilde{\mathbf{Z}}_N^* \tilde{\mathbf{Z}}_N^{*'}) \boldsymbol{\nu}_2,
\end{aligned}$$

where

$$\boldsymbol{\nu}_1 = \begin{pmatrix} c_1 \\ -2 \\ -4\tilde{\mathbf{\Theta}} \end{pmatrix}, \quad \boldsymbol{\nu}_2 = \begin{pmatrix} c_2 \\ -2 \\ -4\tilde{\mathbf{\Theta}}_2 \end{pmatrix}.$$

Applying the Central Limit Theorem to (C.1) and (C.2), we obtain Theorem 4.1. \square

C.2 The proof of Corollary 4.2

To establish Corollary 4.2, we need to calculate ν and σ^2 in (4.8) and (4.9), respectively, for the null case in which the population is $N_{p+q}(\mathbf{0}, \mathbf{I}_{p+q})$. In this case, $\mathbf{X} \sim N_p(\mathbf{0}, \mathbf{I}_p)$, $\mathbf{Y} \sim N_q(\mathbf{0}, \mathbf{I}_q)$ and \mathbf{X} and \mathbf{Y} are independent. Hence, $\|\mathbf{X}\|^2 \sim \chi_p^2$, $\|\mathbf{Y}\|^2 \sim \chi_q^2$, and $\|\mathbf{Z}\|^2 = \|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2 \sim \chi_{p+q}^2$. Therefore, $E(\|\mathbf{Y}\|^4) = q(q+2)$, $E(\|\mathbf{Z}\|^4) = (p+q)(p+q+2)$, and by substituting these results into (4.8), we obtain (4.11).

Next, we calculate σ^2 . Since $\mathbf{Z} \stackrel{\mathcal{L}}{=} -\mathbf{Z}$, it follows from a change of sign that $E(\|\mathbf{Z}\|^4 \mathbf{Z}') = E(\mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z} \mathbf{Z}') = \mathbf{0}$. Similarly, $E(\|\mathbf{Y}\|^4 \mathbf{Y}') = E(\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y} \mathbf{Y}') = \mathbf{0}$. Therefore, (4.9) reduces to

$$\begin{aligned} \sigma^2 &= \frac{c_1^2}{\tau} \text{Var}(\|\mathbf{Z}\|^4) + 4\tau \text{Var}(\mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) + 16\tau \tilde{\mathbf{\Theta}}' \tilde{\mathbf{\Theta}} - 4c_1 \text{Cov}(\|\mathbf{Z}\|^4, \mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) \\ &\quad + \frac{c_2^2}{\bar{\tau}} \text{Var}(\|\mathbf{Y}\|^4) + 4\bar{\tau} \text{Var}(\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) + 16\bar{\tau} \tilde{\mathbf{\Theta}}_2' \tilde{\mathbf{\Theta}}_2 - 4c_2 \text{Cov}(\|\mathbf{Y}\|^4, \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}). \end{aligned}$$

Note that

$$\begin{aligned} \text{Var}(\|\mathbf{Y}\|^4) &= E(\|\mathbf{Y}\|^8) - (E\|\mathbf{Y}\|^4)^2 \\ &= 8q(q+2)(q+3) \end{aligned}$$

and, similarly, $\text{Var}(\|\mathbf{Z}\|^4) = 8(p+q)(p+q+2)(p+q+3)$.

Denote the eigenvalues of $\tilde{\mathbf{\Xi}}_{22}$ by $\lambda_j(\tilde{\mathbf{\Xi}}_{22})$, $j = 1, \dots, q$. The distribution of \mathbf{Y} being orthogonally invariant, we apply an orthogonal transformation to deduce that $\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y} \stackrel{\mathcal{L}}{=} \sum_{j=1}^q \lambda_j(\tilde{\mathbf{\Xi}}_{22}) y_j^2$, where y_1, \dots, y_q are the independent $N(0, 1)$ -distributed components of \mathbf{Y} . Then we obtain $\text{Var}(\mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) = 2 \text{tr}(\tilde{\mathbf{\Xi}}_{22}^2)$ and $\text{Var}(\mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) = 2 \text{tr}(\tilde{\mathbf{\Xi}}^2)$.

As for the covariance terms, we again apply an orthogonal transformation to obtain

$$\begin{aligned} \text{Cov}(\|\mathbf{Y}\|^4, \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) &= \text{Cov}(\|\mathbf{Y}\|^4, \sum_{j=1}^q \lambda_j(\tilde{\mathbf{\Xi}}_{22}) y_j^2) \\ &= \sum_{j=1}^q \lambda_j(\tilde{\mathbf{\Xi}}_{22}) \text{Cov}(\|\mathbf{Y}\|^4, y_j^2) \\ &= \text{tr}(\tilde{\mathbf{\Xi}}_{22}) \text{Cov}(\|\mathbf{Y}\|^4, y_1^2), \end{aligned}$$

where we have also used the exchangeability of y_1, \dots, y_q to deduce that $\text{Cov}(\|\mathbf{Y}\|^4, y_j^2) = \text{Cov}(\|\mathbf{Y}\|^4, y_1^2)$ for all $j = 1, \dots, q$. Since $\|\mathbf{Y}\|^2 \sim \chi_q^2$ then

$$\begin{aligned} \text{Cov}(\|\mathbf{Y}\|^4, y_1^2) &= q^{-1} \sum_{j=1}^q \text{Cov}(\|\mathbf{Y}\|^4, y_j^2) \\ &= q^{-1} \text{Cov}(\|\mathbf{Y}\|^4, \sum_{j=1}^q y_j^2) \\ &= q^{-1} \text{Cov}(\|\mathbf{Y}\|^4, \|\mathbf{Y}\|^2) = 4(q+2). \end{aligned}$$

Hence, $\text{Cov}(\|\mathbf{Y}\|^4, \mathbf{Y}' \tilde{\mathbf{\Xi}}_{22} \mathbf{Y}) = 4(q+2) \text{tr}(\tilde{\mathbf{\Xi}}_{22})$ and, similarly, $\text{Cov}(\|\mathbf{Z}\|^4, \mathbf{Z}' \tilde{\mathbf{\Xi}} \mathbf{Z}) = 4(p+q+2) \text{tr}(\tilde{\mathbf{\Xi}})$.

Collecting together these results, we obtain

$$\begin{aligned} \sigma^2 &= \frac{c_1^2}{\tau} 8(p+q)(p+q+2)(p+q+3) + 8\tau \text{tr}(\tilde{\mathbf{\Xi}}^2) + 16\tau \tilde{\mathbf{\Theta}}' \tilde{\mathbf{\Theta}} - 16c_1(p+q+2) \text{tr}(\tilde{\mathbf{\Xi}}) \\ &\quad + \frac{c_2^2}{\bar{\tau}} 8q(q+2)(q+3) + 8\bar{\tau} \text{tr}(\tilde{\mathbf{\Xi}}_{22}^2) + 16\bar{\tau} \tilde{\mathbf{\Theta}}_2' \tilde{\mathbf{\Theta}}_2 - 16c_2(q+2) \text{tr}(\tilde{\mathbf{\Xi}}_{22}). \end{aligned}$$

Finally, we calculate $\tilde{\Xi}$, $\tilde{\Theta}$, $\tilde{\Xi}_{22}$, and $\tilde{\Theta}_2$. By (4.2),

$$\begin{aligned} \begin{pmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{21} & \Xi_{22} \end{pmatrix} &= \begin{pmatrix} E(\|Z\|^2 \mathbf{X} \mathbf{X}') & E(\|Z\|^2 \mathbf{X} \mathbf{Y}') \\ E(\|Z\|^2 \mathbf{Y} \mathbf{X}') & E(\|Z\|^2 \mathbf{Y} \mathbf{Y}') \end{pmatrix} \\ &= \begin{pmatrix} E(\|Z\|^2 \mathbf{X} \mathbf{X}') & \mathbf{0} \\ \mathbf{0} & E(\|Z\|^2 \mathbf{Y} \mathbf{Y}') \end{pmatrix}. \end{aligned} \quad (\text{C.5})$$

By (4.4),

$$\tilde{\Xi} = \begin{pmatrix} c_1 \Xi_{11} & c_1 \Xi_{12} \\ c_1 \Xi_{21} & c_1 \tau \Xi_{22} + c_2 \bar{\tau} \Xi^* \end{pmatrix} = \begin{pmatrix} c_1 \Xi_{11} & \mathbf{0} \\ \mathbf{0} & c_1 \tau \Xi_{22} + c_2 \bar{\tau} \Xi^* \end{pmatrix}. \quad (\text{C.6})$$

By (4.3),

$$\begin{aligned} \Xi^* = E((\mathbf{Y} \mathbf{Y}')^2) &= E(\mathbf{Y} (\mathbf{Y}' \mathbf{Y}) \mathbf{Y}') \\ &= E(\|\mathbf{Y}\|^2 \mathbf{Y} \mathbf{Y}') = (E(\|\mathbf{Y}\|^2 Y_i Y_j)). \end{aligned}$$

If $i \neq j$ then, by the exchangeability of Y_1, \dots, Y_q , we have

$$\begin{aligned} E(\|\mathbf{Y}\|^2 Y_i Y_j) &= E(\|\mathbf{Y}\|^2 Y_1 Y_2) \\ &= E((Y_1^2 + Y_2^2 + Y_3^2 + \dots + Y_q^2) Y_1 Y_2) = 0, \end{aligned}$$

the last equality following from the mutual independence of Y_1, \dots, Y_q and the fact that each has mean 0. If $i = j$ then by exchangeability,

$$\begin{aligned} E(\|\mathbf{Y}\|^2 Y_i Y_j) &= E(\|\mathbf{Y}\|^2 Y_j^2) \\ &= \frac{1}{q} E(\|\mathbf{Y}\|^2 \cdot \sum_{j=1}^q Y_j^2) = \frac{1}{q} E(\|\mathbf{Y}\|^4) = q + 2; \end{aligned}$$

therefore, $\Xi^* = (q + 2) \mathbf{I}_q$. Next, by (C.5),

$$\begin{aligned} \Xi_{22} = E(\|Z\|^2 \mathbf{Y} \mathbf{Y}') &= E((\|X\|^2 + \|\mathbf{Y}\|^2) \mathbf{Y} \mathbf{Y}') \\ &= E(\|X\|^2) E(\mathbf{Y} \mathbf{Y}') + E(\|\mathbf{Y}\|^2 \mathbf{Y} \mathbf{Y}') = (p + q + 2) \mathbf{I}_q, \end{aligned}$$

and by interchanging the roles of \mathbf{X} and \mathbf{Y} in this latter calculation, we obtain $\Xi_{11} = E(\|Z\|^2 \mathbf{X} \mathbf{X}') = (p + q + 2) \mathbf{I}_p$. Inserting these results at (C.6), we obtain

$$\tilde{\Xi} = \begin{pmatrix} c_1(p + q + 2) \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & (c_1 \tau(p + q + 2) + c_2 \bar{\tau}(q + 2)) \mathbf{I}_q \end{pmatrix}$$

so that $\tilde{\Xi}_{22} = (c_1 \tau(p + q + 2) + c_2 \bar{\tau}(q + 2)) \mathbf{I}_q$.

As for $\tilde{\Theta}$ and $\tilde{\Theta}_2$, it follows from (4.5) and (4.6) that

$$\tilde{\Theta} \equiv \begin{pmatrix} \tilde{\Theta}_1 \\ \tilde{\Theta}_2 \end{pmatrix} = \begin{pmatrix} c_1 \Theta_1 \\ c_1 \tau \Theta_2 + c_2 \bar{\tau} \Theta^* \end{pmatrix}$$

where

$$\Theta = \begin{pmatrix} \Theta_1 \\ \Theta_2 \end{pmatrix} = \begin{pmatrix} E(\|Z\|^2 \mathbf{X}) \\ E(\|Z\|^2 \mathbf{Y}) \end{pmatrix} = \mathbf{0}.$$

By (4.3), $\Theta^* = E(\|\mathbf{Y}\|^2 \mathbf{Y}) = \mathbf{0}$; hence, $\tilde{\Theta} = \mathbf{0}$. Inserting these results in (4.9), we obtain (4.12). \square